

# Weekly Report

23<sup>rd</sup>, Oct. - 5<sup>th</sup>, Nov. 2017

## Done:

1. Finished a report on recent surveyed papers on emergency management. [see below]
2. Discuss my ideas with another member in the lab who may join in the emergency project, Annie. She found a mapping study which covers more paper relevant to infovis in emergency. However, the mapping study do not discuss any of the details of paper (as it is a mapping study), so I need to read some of them in the later days.
3. Revisited of D3 data binding on SVG and drawing and line chart and scatterplot, and coded with one screencast video.

## To Do:

1. Prof. Ma is said to be back, and we need to talk about the project. We need to narrow down the topic.
2. Learn mapping D3 course.

## Paper reading

## Relevant Papers

### Emergency-related system (no visualization, user-interface)

***MADIS: A Multimedia-Aided Disaster Information Integration System for Emergency Management***

**Platform:** Mobile Device iPad

**User:** experts [emergency management personnel assessment of a disaster situation and performing efficient and timely responses]

**Data Source:** Image, textual reports

**Method:** get images taken at disaster scenes along with descriptive info → ML classification (relate images to reports) → user feedback (on images) → ML update

**Comment:** Incorporate User Feedback, update classification results

***Firemap: A Dynamic Data-Driven Predictive Wildfire Modeling and Visualization Environment***

**Platform:** Web

**User:** experts

**Data Source:** all kinds of data (weather, thermal, historical, air quality, images, satellite, ..., corporate with experts from official dept.)

**Method:** UNKNOWN prediction method, only different layers of data are shown

**Comment:** Fusion of multiple highly domain-specific data

More systems can be found at ***Review of wildfire detection using social media***

## **Visual Analysis System**

### ***A Visual Analytics Framework for Microblog Data Analysis at Multiple Scales of Aggregation*** 16' Abish

**Platform:** Web

**User:** Casual Experts (analysts in emergency management, not experts in data science)

**Data source:** Twitter data, map

**Method:** Categorize tweets → show trends of categories → Contextual preserving vis of clusters, categories on the map

**Comment:** support both retrospective and real-time exploration of streaming microblog; smooth transition between different scales (zoom-in and out)

== **POLICE related method** === periodic feature of data [traffic, fire can be periodic]==

### ***Proactive Spatiotemporal Resource Allocation and Predictive Visual Analytics for Community Policing and Law Enforcement*** (Also: A Mobile Visual Analytics Approach for Law Enforcement Situation Awareness) 14' Abish

**Platform:** Desktop

**User:** casual experts

**Data source:** Criminal, Traffic, and Civil (CTC) incident law enforcement datasets

**Method:** create natural scale templates (enable users to focus on appropriate geospatial and temporal resolution levels) → based on historical records, generating time series signal → using STL to predict  $N$  time steps for each geo unit → Visualization

**Comment:** hourly, day-of-week, hourly; from whole picture to user-specific area prediction. Domain knowledge is important; no involving domain knowledge to update model

### ***A Correlative Analysis Process in a Visual Analytics Environment*** 12' Abish

**Platform:** Desktop

**User:** experts

**Data source:** criminal, traffic and civil (CTC) datasets

**Method:** Correlation of different categories of criminals along time, with different lag/lead, entire area/user specified, and the distribution on the map. Considering the periodic feature, visualize by day of a week, hour of a day.

**Comment:** Different ways of correlation computation on different area. Correlation can be 1v1, many v 1. Different correlation computation methods. Potentially extend to recommendation.

### ***Spatiotemporal Social Media Analytics for Abnormal Event Detection and Examination using Seasonal-Trend Decomposition*** 12' Chae

**Platform:**

**User:** experts

**Data source:** Twitter, map, maybe other social media source

**Method:** Evaluate topic anomalies → show tweets on map, show tweets content, show anomaly

along with time

**Comment:** Cross validation with other social data source

***Public behavior response analysis in disaster events utilizing visual analytics of microblog data 14' Chae***

**Platform:** Desktop

**User:** Experts

**Data source:** Twitter with GPS, map, infrastructure data

**Method:** Spatial distribution of tweets comparison (along time), NLP monitor topic anomaly of tweet topics

**Comment:**

***Whisper: Tracing the Spatiotemporal Process of Information Diffusion in Real Time***  
***#FluxFlow: Visual Analysis of Anomalous Information Spreading on Social Media***

**Platform:** Desktop

**User:** Not necessarily experts

**Data Source:** Twitter

**Comment:** quite delicate visual design

**Emergency detection methods (framework, pipeline)**

***Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors***

**Data Source:** twitter data with geotags

**Method:** Semantic Analysis → temporal model (occurrence probability), spatial model (location estimation) → send alert

**Comment:** 1<sup>st</sup> proposes the concept **Social Sensor** No information diffusion (i.i.d.), more statistical models

***The Use of Social Media within the Global Disaster Alert and Coordination System (GDACS)***

**Platform:** Mobile, Web

**User:** Mostly rely on experts

**Data source:** Twitter data

**Method:** Monitoring the abnormal pattern of #tweets contain certain keywords

**Comment:** Information dissemination, monitoring, exchange

***Twitter earthquake detection: earthquake monitoring in a social world***

Similar with the above one + location extraction

### ***General NOAA Oil Modeling Environment (GNOME)***

#### ***Human Sensor Networks for Improved Modeling of Natural Disasters***

**Data:** Flickr (always with images)

**Method:** Retrieve geo information from image EXIF, locate on the map, and conduct comparisons between these and model run outputs

**Comment:** Use retrieved info as a boundary condition (or parameter) in the model

#### ***On Truth Discovery in Social Sensing: A Maximum Likelihood Estimation Approach***

**Data:**  $N$  places to be observed anything wrong (measurement: wrong-1 and OK-0),  $M$  participants. Each participant reports some of place are 1. *DO NOT KNOW* the source reliability and observation correctness.

**Goal:** The correctness of each measurement

**Method:** EM algorithm

**Comment:** without prior knowledge, not in a heuristic manner, 1<sup>st</sup> optimal solution

#### ***“OMG, from here, I can see the flames!”: a use case of mining Location Based Social Networks to acquire spatiotemporal data on forest fires***

**Data:** Twitter data during a forest fire event

**Goal:** retrospective case study

**Comment:** Some insights: (1) Geolocation can be retrieved from both GeoTag & Text (Maybe also IP address if possible); (2) Tweets can be posted by different types of users (citizen, media, ...)

#### ***(Rumor) Prominent Features of Rumor Propagation in Online Social Media***

**Data:** Twitter

**Method:** temporal (epidemic model-like) + structural (graph) + linguistic (NLP-like) → SVM/decision tree/random forest

**Comments:** A comprehensive method.

## **Ideas and notes**

We need to narrow down the topic at first. Discussion with Dana about their requirements, the end users.

### **Data pre-processing**

- Fire historical data/reports
  - Learn the form of reports, the problems in the reports
  - Data cleaning
  - Retrieve key information (when where what)

- Social media data
  - Data quality investigation
  - Geotag or location available?
  - Linguistic analysis, (when where what) turn into easy-to-use format
- Weather (wind speed, direction, humidity...)
  - Available in what time granularity
  - In what space granularity
- Traffic
  - Available in what time granularity
  - In what space granularity

### Goals & potential usage scenarios

The goal is to raise *situational awareness*. Monitoring and Prediction.

- Some usage scenarios:
  - Place A on fire → based on DATA, predict whether the fire will spread to Yolo county?
  - Place A on fire → residents from A will evacuate to Yolo, and the influence on Yolo traffic?
  - Yolo on fire → how do we evacuate based on the traffic, weather condition, real-time fire condition | how to allocate resources on the evacuation route based on more information like the residence info, infrastructure info.

### Ideas

As far as I know,

- Prior papers seldom mention the **uncertainty issues** in the data. In fact, there are many uncertainties can be found:
  - Social media data -- reliability of tweets content, the geolocation (*noise* in social sensor)
  - The reports are said to have many different formats, with time delays
  - The fire area with delays
  - The traffic data also not that accurate (i.e., may not be real time, also with delays)
- Prior relevant vis papers (emergency management, predictive policing) EITHER only utilize social media data OR only use event data (like criminal records, similar to fire records). When they use social media, no structural info (graph properties of social network are utilized)
- Most of their prediction methods DO NOT involve much experts in the loop (majority of these methods, expert knowledge is utilized as part of the model input)

- We may develop a recommendation model. Incorporate weather data, historical records, domain knowledge. Users can visually predict based on visualization results, with the help of recommendation OR help tune the recommendation result with domain knowledge. User interactions, visualization updates are included.

Use picnic day as case: In 2010, the Sacramento Bee reported that campus and community leaders are considering canceling UC Davis' annual Picnic Day due to the high number of arrests and mayhem during Picnic Day 2010. If data are available, a potential good case.

## Other potential topics

**Evacuation Routine Plan**

**Rumor dissemination & control**

...MORE

## Other resources

Satellite images for California: [Open California by Planet Labs](#)

...MORE

### 安排表

内容	DDL	Milestone
迭代式讨论研究主题	应该近期能大致推定	做完概要，继续捡漏
D3 学习以及框架	A.S.A.P	基本知识大致又过了一遍，对于原理性的东西有了新的理解。后面要开始接触实际画图，不同视图关联等。